Extraction of arbitrarily-shaped objects using stochastic multiple birth-and-death dynamics and active contours

Maria Kulikova, Ian Jermyn, Xavier Descombes, Elena Zhizhina, Josiane Zerubia

Extraction of arbitrarily-shaped objects using stochastic multiple birth-and-death dynamics and active contours

Maria S. Kulikova[a], Ian H. Jermyn[a], Xavier Descombes[a], Elena Zhizhina[b], and Josiane Zerubia[a]

[a]EPI Ariana, CR INRIA Sophia Antipolis Méditerranée, 2004 route des Lucioles, B. P. 93, 06902, Sophia Antipolis, France;
[b]Dobrushin Laboratory of Mathematics, Institute of Information Transmission Problems (IITP), Bolshoy Karetny per. 19, 127994 Moscow, Russia.

ABSTRACT

We extend the marked point process models that have been used for object extraction from images to arbitrarily shaped objects, without greatly increasing the computational complexity of sampling and estimation. The approach can be viewed as an extension of the active contour methodology to an *a priori* unknown number of objects. Sampling and estimation are based on a stochastic birth-and-death process defined in a space of multiple, arbitrarily shaped objects, where the objects are defined by the image data and prior information. The performance of the approach is demonstrated via experimental results on synthetic and real data.

Keywords: object extraction, marked point process, active contour, multiple birth-and-death dynamics

1. INTRODUCTION

The resolution of optical satellite and aerial images is continually increasing, with resolutions now ranging from several decimetres down to several centimetres. At these resolutions, the geometry of objects at human scale is clearly visible, and needs to be taken into account in analysing the images. Stochastic point process models of multiple objects and their configurations are more easily able to include this type of geometrical information than classical approaches such as Markov random fields. A probability density is defined on the space of configurations of multiple objects that depends on the objects' relation to the data, and the individual and joint configuration of the objects. The optimal configuration of objects is then estimated, usually using Maximum A Posteriori (MAP) estimation. In previous work, the objects involved have been represented using simple geometrical shapes, *e.g.* discs, ellipses, or rectangles. The resulting models have been applied to the extraction of different types of object from remote sensing images, *e.g.* road networks,[9] buildings,[11] and trees,[13] but the simplified nature of the individual objects limits the geometrical precision that can be achieved.

The aim of this paper is to lift this restriction without increasing the dimension of the single-object space (*e.g.* 1 for discs, 3 for ellipses and rectangles), since this would greatly increase the computational complexity of sampling and estimation. The single-object space considered is thus still of small dimension, but the possible individual objects are determined not *a priori*, but by the image data and a single-object version of the model. As a result, they can be arbitrary closed curves. Once the single-object space is defined, a Gibbs energy, and hence a probability distribution, are defined on the configuration space of an arbitrary number of single objects. The energy includes single-object terms that relate each object to the image and control boundary smoothness, and which were used in the definition of the single-object space, and an interaction term that penalizes object overlap. To find the MAP estimate using the full energy, we sample using a multiple birth-and-death process embedded in an annealing scheme.[4]

The use of a birth-and-death process allows the number of objects to be unknown *a priori*. Our work can thus be thought of as an extension of the active contour methodology[7] to an *a priori* unknown number of objects. A great deal of work has been done within the active contour framework using the distance-function level set representation.[1,3,10,12,14] This representation allows arbitrary topology, *i.e.* an arbitrary number of contours, at

Contact: M. S. Kulikova (maria.kulikova@inria.fr).

least in principle. However, it is a representation of a region with arbitrary topology, not an arbitrary number of distinct objects: overlapping objects, for example, cannot be represented. Second, the algorithms used are forms of deterministic gradient descent, meaning that the result may be very dependent on the initial configuration, particularly since arbitrary topology changes are not possible during the algorithm. Cremers *et al.* [2] treat the case in which there is a number of distinct classes of object in the image by segmenting the image into connected components each of which corresponds to one class. The representation is by distance-function level sets, however, so that overlapping objects are not allowed, while the way in which prior information is included means that only one object can be found in each connected component. There are many tracking methods that use stochastic algorithms, *e.g.* ,[5,8,16] but those that deal with multiple objects mostly do so by using the distance-function level set representation. Storvik[15] uses an MCMC algorithm to minimize an active contour energy, but considers only simply-connected objects; the algorithm makes only local changes to the contour at each iteration. Juan *et al.* [6] use stochastic PDE techniques for optimization, but again the stochastic element is limited to small changes to the contour. In contrast, in this paper, although the possible forms of single-object are limited by adaptation to the data, multiple objects can be created and destroyed at each iteration.

In section 2, we describe the single-object space and the single-object terms in the energy. In section 3, we describe the multiple-object space and the full energy, as well as the sampling algorithm. In section 4, we describe experimental results validating the algorithm, and in section 5, we conclude.

2. SINGLE OBJECTS

We take as the single-object space a subset of the space of all possible arbitrarily-shaped objects. An object is represented by its boundary, a closed planar curve $\gamma : [0, 2\pi] \to V \subset \mathbb{R}^2$ lying in the image domain V. Given an energy functional E_d defined on a space Γ of such curves, and an initial curve $\gamma \in \Gamma$, we can perform gradient descent to obtain a curve, $\hat{\gamma} \in \Gamma$ given by a local minimum of E_d; this defines the map $\hat{\cdot} : \Gamma \to \Gamma$. Now denote by $\mathcal{C}$ the set of parameterized circles lying in the image domain, with radii in $[r_{\min}, r_{\max}]$ and arc-length parameterization. The single-object space we consider is $\Gamma_o = \hat{\mathcal{C}}$. The objects thus obtained from the circles are locally adapted to the data, since the energy functional depends on the image as well as curve geometry. The dimension of the single-object space is still small, however: if we fix the centre of the circle, it is at most one-dimensional; in practice, it is very likely to be a discrete set.

2.1 Single-object energy

The energy E_d used to construct the space Γ_o is a weighted sum of three terms:

$$E_d(\gamma) = \frac{\lambda}{2} \int_{[0,2\pi]} dt\, |\dot{\gamma}(t)|^2 + \lambda_i \int_{[0,2\pi]} dt\, n(t) \cdot \nabla I(\gamma(t)) + \alpha \int_{R(\gamma)} d^2x\, (G(x) - \bar{G}(x)) \,,$$

where $\dot{\cdot}$ denotes a derivative; n is the (unnormalized) outward normal to the curve; I is the image; $G(x) = \frac{(I(x)-\mu)^2}{2\sigma^2}$; $\bar{G}(x) = \frac{(I(x)-\bar{\mu})^2}{2\bar{\sigma}^2}$; and $R(\gamma)$ is the region corresponding to the boundary γ. The first term favours smooth boundaries and uniform parameterization. The second term favours object boundaries orthogonal to the image gradient. The third term arises from a Gaussian image model with means μ, $\bar{\mu}$ and variances σ, $\bar{\sigma}$ for the interior and exterior of the objects respectively. These parameters are learned from pieces of the image.

The algorithm makes uses of the functional derivative of E_d:

$$\frac{\delta E_d}{\delta \gamma(t)} = -\frac{\lambda}{2}\ddot{\gamma}(t) - n(t)\big[\lambda_i \nabla^2 I(\gamma(t)) + (G(\gamma(t)) - \bar{G}(\gamma(t)))\big] \,. \tag{1}$$

Figures 1 and 2 show examples of gradient descent from initial circles using the energy E_d, showing how the points of Γ_o, generated as they are by the map $\hat{\cdot}$, are locally adapted to the image.

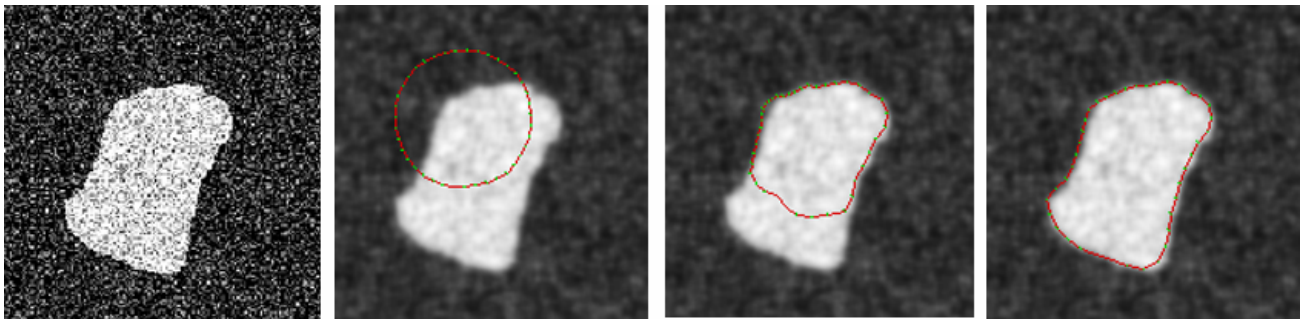

Figure 1: Synthetic image example: gradient descent driven by E_d. Left: binary image with added noise. The other three images show the contour evolution. The background has been blurred for display purposes.

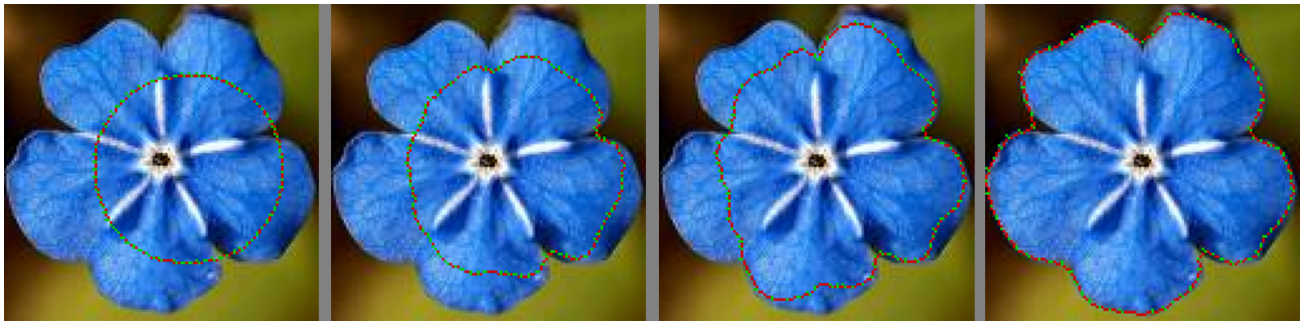

Figure 2: Real image example: gradient descent driven by E_d. Only one band of the colour image was used.

3. MULTIPLE OBJECTS: MODEL AND ALGORITHM

The multiple-object space is the set of configurations of zero or more objects:

$$\Omega_{\Gamma_o} = \bigcup_{n=0}^{\infty} \left[\Gamma_o^n / S_n \right] , \tag{2}$$

where S_n indicates the symmetric group of n elements acting on the components of the product. Note that set union and difference are defined for appropriate pairs of elements of Ω_{Γ_o}. The map $\hat{\cdot}$ extends to a map from $\Omega_{\mathcal{C}}$ (the space of configurations of circles from $\mathcal{C}$) to Ω_{Γ_o}. We denote by ω and $\hat{\omega}$ elements of $\Omega_{\mathcal{C}}$ and Ω_{Γ_o} respectively.

3.1 Energy

Denote by H a real function on $\Omega_{\mathcal{C}}$ that is bounded below. We then define the Gibbs distribution μ_β in terms of a density $p(\omega) = \frac{d\mu_\beta}{d\lambda}(\omega)$ with respect to Lebesgue-Poisson measure λ on $\Omega_{\mathcal{C}}$:

$$p(\omega) = \frac{z^{|\omega|}}{Z_\beta} \exp\{-\beta H(\omega)\} , \tag{3}$$

with parameters $\beta > 0$, $z > 0$ and a normalizing factor Z_β:

$$Z_\beta = \int_{\Omega_{\mathcal{C}}} d\lambda(\omega)\, z^{|\omega|}\, \exp\{-\beta H(\omega)\} .$$

The energy H is a sum of two terms:

$$H(\omega) = c_0 \sum_i H_1(\omega_i) + \sum_{i \neq j} H_2(\omega_i, \omega_j) ,$$

where $c_0 \in \mathbb{R}$ and the $\{\omega_i\}$ are the elements of $\omega \in \Omega_{\mathcal{C}}$. The term H_1 is

$$H_1(\omega_i) = E_d(\hat{\omega}_i) \ .$$

The term H_2 introduces an interaction between objects that penalizes overlaps:

$$H_2(\omega_i, \omega_j) = \frac{A(R(\hat{\omega}_i) \cap R(\hat{\omega}_j))}{\min(A(R(\hat{\omega}_i)), A(R(\hat{\omega}_j)))} + \delta_\epsilon(\omega_i, \omega_j) \ ,$$

where A is the area functional and δ_ϵ is a hard-core repulsion between two components of ω, thereby eliminating some small region around the diagonal in $\Omega_{\mathcal{C}}$. This prevents the "condensation" of an infinite number of objects in the lowest energy single-object configuration. Note that this is not guaranteed to eliminate the diagonal in Ω_{Γ_o}. To achieve that, the interaction term must be sufficiently strong with respect to the single-object terms.

Apart from the hard-core repulsion, the energy H is constant on the equivalence classes defined by the map $\hat{\cdot}$, *i.e.* it is the same for all $\omega \in \Omega_{\mathcal{C}}$ that project to the same $\hat{\omega} \in \Omega_{\Gamma_o}$. However, since the hard-core repulsion is equivalent to eliminating some elements close to the diagonal in $\Omega_{\mathcal{C}}$, it is equivalent to the elimination of some elements of Ω_{Γ_o}. Thus the relevant multiple-object space is not quite given by (2), but rather by this space minus the eliminated elements. We will use the same symbol Ω_{Γ_o} for both. (Algorithmically, we will work in $\Omega_{\mathcal{C}}$, so the difference will be transparent.) With this redefinition, the energy H is well-defined as an energy on Ω_{Γ_o}.*

3.2 Sampling and estimation

To obtain a MAP estimate of the configuration of the objects in the image, we sample from the probability distribution μ_β using a Markov chain in $\Omega_{\mathcal{C}}$ in conjunction with an annealing scheme. The chain has a discrete-time multiple birth-and-death dynamics that converges to a continuous process as the time-step tends to zero. The reader may refer to[4] for details of the convergence properties.

We restrict $\mathcal{C}$ to be the set of circles centred at the image pixels, with radii in the range $[r_{\min}, r_{\max}]$. All curves are represented in the algorithm by points in $\mathbb{R}^2$ defined to correspond to discrete parameter values $t_n = 2\pi n/N$ for $n \in \{0, \ldots, (N-1)\}$. The circles in $\mathcal{C}$ are arc-length parameterized, and when discretized are represented by equally spaced points.

At each birth step, circles sampled from a uniform Lebesgue-Poisson process with an intensity z that is dependent on the discrete-time step δ, but that is independent of the current temperature and the energy of the current configuration, are added to the current configuration. At each death step, each object is removed from the current configuration with a probability that is a function of

$$d_\beta(\omega_i, \omega) = e^{-\beta H(\omega_i, \omega)} \ , \tag{4}$$

where β is the current (inverse) temperature, and $H(\omega_i, \omega) = H(\omega - \omega_i) - H(\omega)$ is the energy difference between the configurations with and without the object. In more detail:

1. **Initialization**

   Discrete time-step $\delta = \delta_0$; inverse temperature $\beta = \beta_0$; Poisson intensity z_0; radius range $[r_{\min}, r_{\max}]$; parameters in E_d;

2. **Birth**

   (a) Sample a configuration of circles (with radii uniformly sampled from $[r_{\min}, r_{\max}]$) from the Lebesgue-Poisson distribution with intensity $z = \delta z_0$, with the addition of a hard core repulsion δ_ϵ with ϵ equal to one pixel, and add them to the existing configuration, producing configuration $\omega \in \Omega_{\mathcal{C}}$;

*Note, however, that this is not the same as pushing forward μ_β from $\Omega_{\mathcal{C}}$ to Ω_{Γ_o} and then taking a MAP estimate, since there is no reason to believe that the entropy factor that would be generated by this procedure would be constant on Ω_{Γ_o}.

(b) Evolve every circle in ω using gradient descent, with gradient field given by equation (1), until convergence, producing configuration $\hat{\omega} \in \Omega_{\Gamma_o}$;

3. **Death**

   (a) For computational efficiency, sort the components of the current configuration with respect to their energy $H_1(\omega_i) = E_d(\hat{\omega}_i)$;

   (b) Remove each object ω_i from the current configuration with probability

$$p_d(\omega_i, \omega) = \frac{\delta d_\beta(\omega_i, \omega)}{1 + \delta d_\beta(\omega_i, \omega)} \; ;$$

4. **Termination**

   If all the components added in the birth step and only these, are removed in the following death step, then stop; if not, then decrease the temperature $T = \frac{1}{\beta}$ and time-step δ, and go to the birth step.

Note that this process finds an energy minimizing $\omega \in \Omega_{\mathcal{C}}$, which then projects to $\hat{\omega} \in \Omega_{\Gamma_o}$. The algorithm terminates despite the degeneracy of the energy on $\Omega_{\mathcal{C}}$ (see the discussion at the end of section 3.1) because the birth step always moves the configuration out of its current equivalence class under $\hat{\cdot}$.

4. EXPERIMENTAL RESULTS

Figures 3 and 4 show the results of experiments on a synthetic binary image with added white Gaussian noise, and on one band of a Colour Infrared (CIR) very high resolution aerial image showing tree crowns from vertically above. Note that the initial configuration contains no curves. The first birth step creates a certain number of curves, but the final number is determined automatically by the annealed birth-and-death process.

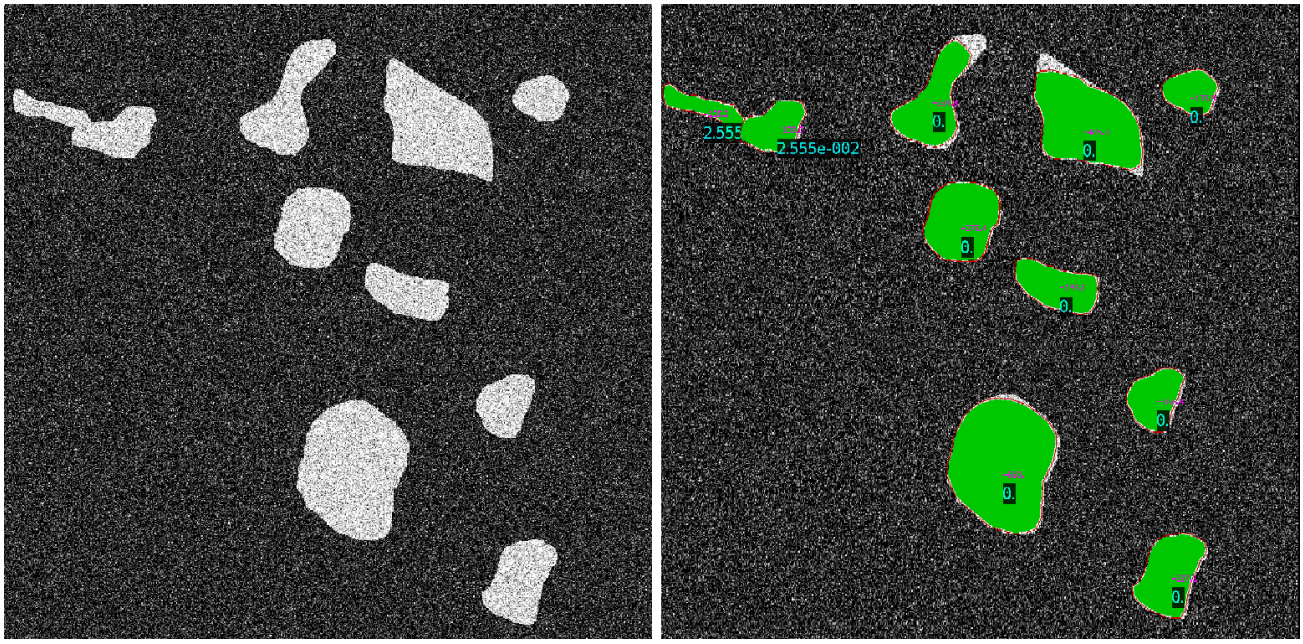

Figure 3: Left: synthetic binary image with added noise. Right: final object configuration. The numbers in the interiors of the curves show the value of H_2 (black background), demonstrating that there is no degeneracy in the solution, and the value of H_1.

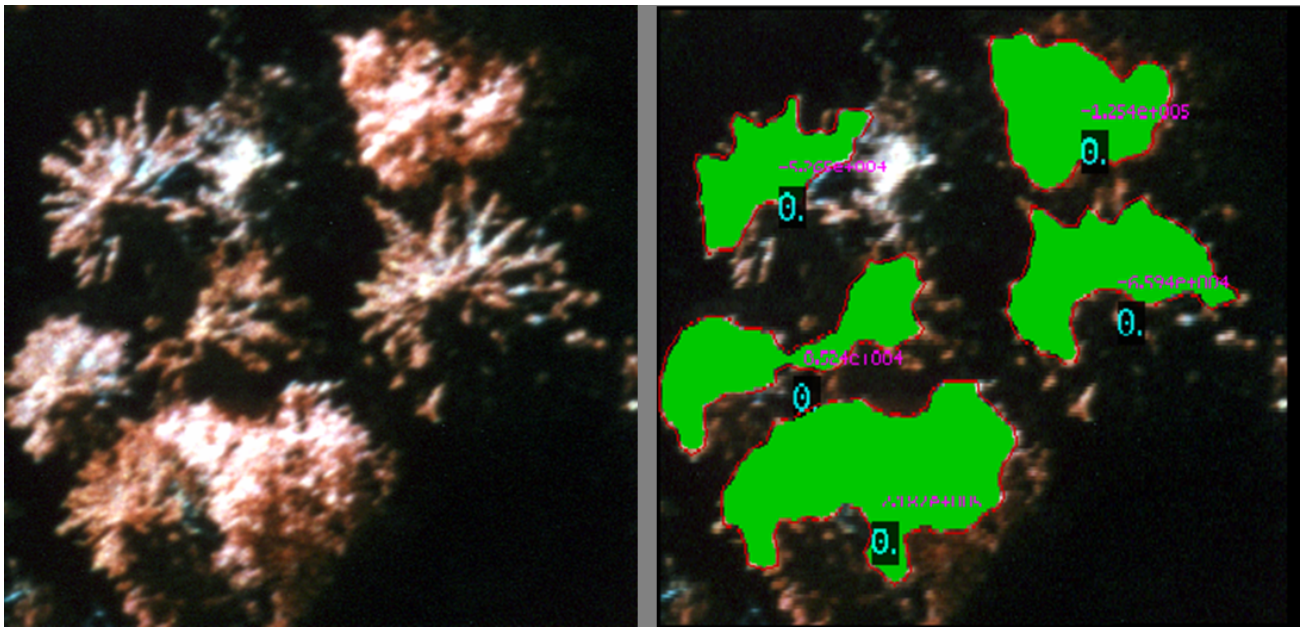

Figure 4: Left: CIR image of tree crowns, ©CBA. Right: final object configuration. The numbers in the interiors of the curves show the value of H_2 (black background), demonstrating that there is no degeneracy in the solution, and the value of H_1.

5. CONCLUSIONS AND PERSPECTIVES

The marked point process models that have been used in the past for object extraction from images have used only simple geometric objects. We have proposed a method for the incorporation of arbitrarily complex shapes, without significantly increasing the computational complexity of sampling and estimation. The work can also be seen as an extension of the active contour methodology to an *a priori* unknown number of objects. The set of potential objects is defined using an energy that incorporates prior information as well as information coming from the data. The model could therefore be extended by including more specific prior information about object shape without significantly increasing the computational load. The birth-and-death algorithm used has the advantage that at every iteration, the current configuration is updated by adding multiple objects independently of the current energy and temperature. This increases computational efficiency.

As the experimental results show, the model allows the accurate delineation of objects with complex shapes. However, this very precision can lead to inaccuracy in the number of extracted objects: two overlapping crowns are detected by a single object in the lower part of figure 4, for example. In order to correct this, one could increase the weight of the smoothing term. This would further penalize object boundary length, and would thereby tend to prevent the extraction of a group of tree crowns by a single object. The drawback is that the resulting single object space would contain only highly smoothed curves, thereby simplifying the shapes of the extracted objects. The longer-term goal of this work is therefore to incorporate more specific prior information into the energy E_d, in order to detect overlapping objects, and eventually, simultaneously to extract and classify several types of objects in scenes of high complexity, *e.g.* tree species in aerial images of dense forest.

ACKNOWLEDGEMENTS

This work was performed within the framework of, and partly financed by, INRIA/IITP/UIIP Associated Team "ODESSA" (http://www-sop.inria.fr/ariana/Projets/Odessa/) and INRIA/FSU Associated Team "SHAPES" (http://www-sop.inria.fr /ariana/Projets/Shapes/). The authors thank the Swedish University of Agricultural Sciences of Uppsala for the CIR data.

REFERENCES

[1] Caselles, V., Kimmel, R., Sapiro, G.: "Geodesic active contours." Int. J. Comp. Vis., **22(1)**: 61–79 (1997)

[2] Cremers, D., Sochen, N., Schnörr, C.: "A multiphase dynamic labeling model for variational recognition-driven image segmentation." Int. J. Comp. Vis., **66(1)**: 67–81 (2006)

[3] Cremers, D., Osher, S.J., Soatto, S.: "Kernel density estimation and intrinsic alignment for shape priors in level set segmentation." Int. J. Comp. Vis., **69(3)**: 335–351 (2006)

[4] Descombes, X., Minlos, R., Zhizhina, E.: "Object extraction using a stochastic birth-and-death dynamics in continuum." J. Math. Imaging Vis.,**33(3)**: 347–359 (2009)

[5] Isard, M., Blake, A.: "Condensation–Conditional Density Propagation for Visual Tracking." Int. J. Comp. Vis., **29(1)**: 5–28 (1998)

[6] Juan, O., Keriven, R., Postelnicu, G.: "Stochastic Motion and the Level Set Method in Computer Vision: Stochastic Active Contours." Int. J. Comp. Vis., **69(1)**: 7–25 (2006)

[7] Kass, M., Wiktkin, A., Terzopolous, D.: "Shakes: Active contours models." Int. J. Comp. Vis., **1(4)**: 321–331 (1998)

[8] Kervann, C., Heitz, H. : "A hierarchical Markov modeling approach for the segmentation and tracking of deformable shapes." Graph. Mod. and Im. Proc., **60**: 173–195 (1998)

[9] Lacoste, C., Descombes, X., Zerubia, J.: "Point processe for unsupervised line network extraction in remote sensing." IEEE Trans. on Patt. Anal. Mach. Intell., **27(10)**: 1568–1579 (2005)

[10] Leventon, M.E., Grimson, W.E.L., Faugeras, O.: "Statistical shape influence in geodesic active contours." In Proc. CVPR, **1**: 316–323 (2000)

[11] Ortner, M., Descombes, X., Zerubia. J.: "Building outline extraction from digital elevation models using marked point processes." Int. J. Comp. Vis., **72(2)**: 107–132 (2007)

[12] Osher, S., Fedkiw, R.: "Level Set Methods and Dynamic Implicit Surfaces." Springer Verlag (2003)

[13] Perrin, G., Descombes, X., Zerubia, J.: "A marked point process model for tree crown extraction in plantation." In Proc. ICIP (2005)

[14] Sethian, J.A.: "Level Set Methods and Fast Marching Methods." Cambridge Univ. Press (1999)

[15] Storvik, G.: "A Bayesian approach to dynamic contours through stochastic sampling and simulated annealing." IEEE Trans. on Patt. Anal. Mach. Intell., **16(10)**: 976–986 (1994)

[16] Rathi, Y., Namrata Vaswani, N., Tannenbaum, A., Yezzi, A.: "Tracking Deforming Objects Using Particle Filtering for Geometric Active Contours." IEEE Trans. on Patt. Anal. Mach. Intell., **29(8)**: 1470–1475 (2007)